Transportation Demand Modeling: Econometric Analysis of SP Experiments Constructed from RP Choices

Kenneth Train
University of California, Berkeley
And
Wesley W. Wilson
University of Oregon

Motivation

- Insufficient variation in RP data: need to augment with SP data
- Standard SP experiments are problematic for analysis of freight shipping:
  - Can be unrealistic since the numerous real-world factors that affect shippers cannot be incorporated
  - Shippers respond to the lack of realism by not answering or answering arbitrarily

SP questions based on RP choice

- What are your alternatives for making shipments?
- Which alternative did you choose for your last shipment?
- Suppose the rate for that alternative rose by 40%. Would you still choose that alternative or would you switch to a different alternative?

SP questions based on RP choice

- Alternatives are the same as in RP setting.
- Attributes of alternatives are the same as in RP setting, except for the one change that is stated.
- Shipper is accustomed to making this kind of choice and can consider all factors, including unobserved factors, that affect their choice.

Similar SP procedures by other researchers

- Pivoted SP experiments of Greene, Hensher, Rose, and others
  - Identify a “recent trip” of respondent
  - Construct SP alternatives by changing attributes of recent trip.
  - Maybe include recent trip as an alternative.
- Leeds Adaptive Stated Preference of Fowkes and Shinghal
  - Create a pivoted SP experiment from a chosen RP alternative
  - Observe choice in this SP experiment
  - Change attributes for next SP experiment based on this choice
  - Continue for numerous SP experiments

Econometric issues

- SP attributes are endogenous since they are based on the RP choice.
- Unobserved factors that affect RP choice can be expected also to affect SP choice.
- We develop an estimation procedure where:
  - Errors in RP setting enter the SP choices
  - SP choices are conditional on RP choice

Survey

- Implemented by the Social and Economic Sciences Research Center at Washington State University
- Shippers located both on and off Columbia/Snake waterway
- 181 observations, predominantly grain shippers

Alternatives

1. Truck to Pasco and barge to Portland
2. Truck to another barge port and barge to Portland
3. Rail to Portland
4. Truck to a rail terminal and rail to Portland
5. Barge to Portland
6. Other.

Survey Information
RP Choices

- Shippers were asked to consider their last shipment.
- Identify the alternatives available to them
  - Rates, transit time, and reliability for each alternative
- Identify which alternative they chose.

Survey Information
SP Choices

- Suppose the rate for your chosen alternative rose by X% and rates for other alternatives remained the same. Would you continue to use that alternative or would you switch to a different alternative? If switch, to what alternative?
- X is randomly chosen from 10, 20, …, 60.
- Similar questions for transit times and reliability.

Fixed Coefficients Specification

Utility from RP alternatives:

$$U_j = \beta x_j + \varepsilon_j.$$

RP choice probability is standard logit:

$$P_i = \frac{e^{\beta x_i}}{\sum e^{\beta x_j}}.$$

$\tilde{x}^i_{jt}$

Stated-Preference choice

- Same alternatives as in RP setting
- Attribute $\tilde{x}^i_{jt}$ for alternative *j* in question *t* based on alternative *i* having been chosen in the RP setting.
- Utility of alternative j in SP setting:

$$W_{jt} = \beta \tilde{x}^i_{jt} + \varepsilon_j + \eta_{jt}.$$

where new error has scale 1/α

Stated-Preference choice

- Equivalently:

$$W_{jt} = \alpha\beta\tilde{x}^i_{jt} + \alpha\varepsilon_j + \eta_{jt}$$

- Choice probability is mixed logit with mixing over conditional distribution of ε:

$$P_{k_t|i} = \int \frac{e^{\alpha\beta\tilde{x}^i_{k_t t} + \alpha\varepsilon_{k_t}}}{\sum_j e^{\alpha\beta\tilde{x}^i_{jt} + \alpha\varepsilon_j}} f\left(\varepsilon \mid i \text{ chosen}\right) \mathrm{d}\varepsilon.$$

Draws from conditional density of ε

Chosen RP alternative:

$\varepsilon_i \sim$ extreme value shifted up by $-ln(P_i)$

Non-chosen RP alternatives:

$\varepsilon_j \sim$ extreme value truncated at $\beta x_i - \beta x_j + \varepsilon_i$

Model on RP data only, fixed coefficients

Variable	Estimate	T-statistic
Rate, in dollars per ton	-0.1252	1.977
Time, in days	-0.0342	1.070
Reliability	0.0322	2.839
Value of time	1 day = 27 cents per ton	
Value of reliability per ton	1 percent = 26 cents	

Model on RP and SP data, fixed coefficients

Variable	Estimate	T-statistic
Rate, in dollars per ton	-0.2086	5.625
Time, in days	-0.1483	6.356
Reliability	0.0282	6.127
Scale of sp error (α)	5.587	3.444
Value of time	1 day = 71 cents per ton	
Value of reliability per ton	1 percent = 14 cents	

Random coefficients

- β is random instead of fixed
- RP choice probability is mixed logit, mixed over distribution of β.
- SP choice probabilities is mixed logit, mixed over distribution of β and over conditional distribution of ε.

Model on RP and SP data, random coefficients

- Time and reliability coefficients: censored normal (with mass at zero)
- Rate coefficient fixed.
- Scale set to $\alpha=10$ after preliminary analysis indicates unboundedly large value.

Model on RP and SP data, random coefficients

Variable	Estimate	T-statistic
Rate	-0.2325	7.610
Time: mean	-0.3031	5.027
Time: std dev	0.2235	3.448
Reliability: mean	0.0367	6.756
Reliability: std dev	0.0170	3.777
Mean value of time	1 day = $1.34 per ton	
Mean value of reliability per ton	1 percent = 16 cents	